

AWARD NUMBER: W81XWH-13-1-0451

TITLE: The Genomic Evolution of Prostate Cancer

PRINCIPAL INVESTIGATOR: David VanderWeele, MD PhD

CONTRACTING ORGANIZATION: University of Chicago. Ejecta. "N" 828559

REPORT DATE: October 2014

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE		<i>Form Approved</i> OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.			
1. REPORT DATE "Qexqdt"2014	2. REPORT TYPE Annual	3. DATES COVERED 52"Ugr 2013 – 4; Sept 2014	
4. TITLE AND SUBTITLE The Genomic Evolution of Prostate Cancer		5a. CONTRACT NUMBER W81XWH-13-1-0451	
		5b. GRANT NUMBER Y : 3ZYJ /35/3/2673	
		5c. PROGRAM ELEMENT NUMBER	
		5d. PROJECT NUMBER	
6. AUTHOR(S) David VanderWeele E-Mail: dvanderw@medicine.bsd.uchicago.edu		5e. TASK NUMBER	
		5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) University of Chicago, The 5801 S Ellis Ave Chicago IL 60637-5418		8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Medical Research and Materiel Command Fort Detrick, Maryland 21702-5012		10. SPONSOR/MONITOR'S ACRONYM(S)	
		11. SPONSOR/MONITOR'S NUMBER(S)	
12. DISTRIBUTION / AVAILABILITY STATEMENT Approved for Public Release; Distribution Unlimited			
13. SUPPLEMENTARY NOTES			

14. ABSTRACT

Prostate cancer is a heterogeneous disease characterized by diverse outcomes. A subset of patients originally diagnosed with low risk disease go on to be diagnosed with high grade disease. It is important to determine to what extent this is due to undersampling on the initial biopsy, and to what extent low grade disease evolves. In addition, multiple genetic alterations are associated with disease evolution in response to therapy. This project aims to characterize evolution of prostate cancer. Completed work used exome sequencing to evaluate the relationship of low grade foci, high grade foci, and metastases in radical prostatectomy specimens. This work demonstrates that coincident low and high grade disease are distantly related, indicating that an early parental clone can give rise to both low grade and high grade disease. Conversely, lymph node metastases are closely related to high grade cancer. Alterations in the TP53 pathway are associated with high grade disease. Aims still under investigation are evaluating the relationship between coincident conventional prostate cancer and the aggressive variant ductal adenocarcinoma, and using circulating tumor cells to evaluate the evolution of castrate resistant metastatic cancer from primary foci.

15. SUBJECT TERMS

Cancer genetics, tumor evolution, tumor heterogeneity, prostate cancer, exome sequencing

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON USAMRMC
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER <i>(include area code)</i>
Unclassified	Unclassified	Unclassified	Unclassified	18	

Standard Form 298
(Rev. 8-98)
Prescribed by ANSI
Std. Z39.18

Table of Contents

Introduction	5
Keywords	5
Overall Project Summary	5
Key Research Accomplishments	9
Conclusion	9
Publications	10
Inventions, Patents and Licenses	10
Reportable Outcomes	10
Other Achievements	10
Opportunities for Training	10
Opportunities for Professional Development	11
References	11
Appendices	11

Introduction

Prostate cancer is a heterogeneous disease characterized by diverse outcomes. A subset of patients originally diagnosed with low risk disease are later diagnosed with high grade disease. Measures taken to reduce these events depend on to what extent this is due to undersampling on the initial biopsy, and to what extent low grade disease evolves. Whereas it is unclear if low grade, indolent cancer evolves into high grade, aggressive disease, select genetic alterations are associated with evolution of advanced disease in response to therapy. This project aims to characterize evolution of prostate cancer in both localized and metastatic disease. One aim seeks to evaluate evolution of localized disease to high grade disease or aggressive variants. The second aim seeks to characterize evolution that occurs in the progression of localized prostate cancer to metastatic disease.

Keywords

Cancer genetics, tumor evolution, tumor heterogeneity, prostate cancer, exome sequencing

Overall Project Summary

The research project proposed within this award was divided into two specific aims, with sub aims further delineated within each aim and specific tasks outlined within the statement of work. This portion of the report will be organized by specific aim and according to the tasks outlined in the SOW.

Aim 1: Investigate the molecular relationship between coincident prostate cancer foci.

Aim 1a: *Test the hypothesis that coincident low and high grade foci represent evolution of a single clone.*

Matched low and high grade foci from prostatectomy specimens underwent exome sequencing to an average depth of 75 million reads (Table 1). Five specimens were sequenced. One specimen was sequenced twice but with low quality data each time, and so data from four specimens were used for evaluation. Sequences were aligned and 301 potential somatic mutations were identified using VarScan and custom scripts. A portion of high confidence somatic mutations identified through VarScan were confirmed via Sanger sequencing and targeted capture and resequencing (Figure 1).

Table 1. Characteristics of sequencing libraries

Patient number	18	18	18	18	14	14	14	6	6	6	6	25	25	2
Histology	GI 6	GI 8	LN	NI	GI 6	GI 8	NI	GI 6	GI 8	LN	NI	GI 6	GI 9	N
Unique mapped reads (millions)	92.9	62.9	53.4	90.6	67.3	78.6	66.8	113.4	75.3	58.8	86.4	69.9	78.3	64
Median fold coverage	63	41	50	59	33	32	31	79	46	51	29	14	27	2
Fraction covered at 10x	0.92	0.89	0.93	0.92	0.87	0.85	0.87	0.93	0.92	0.95	0.79	0.73	0.89	0.7
High confidence mutations	11	23	23		15	45		19	57	60		34	14	
Mutation rate per Mb	0.30	0.63	0.63		0.41	1.23		0.52	1.56	1.64		0.93	0.38	

High confidence mutations include nonsynonymous, synonymous, UTR (untranslated region), and splice junction mutations, and introductions of stop codons. Mutation rate is based on 36.5 Mb of exome sequence targeted for capture.

The results of this aim demonstrate that low grade prostate cancer diverges early from high grade prostate cancer (Figure 2), and that in some cases we were not able to rule out that low and high grade cancer derived from separate clones.

We also found that mutations in the p53 signaling pathway are exclusive to high grade foci, suggesting that these mutations may be a marker of high risk disease. Results from this aim are now published and the manuscript is included as an appendix to this report (VanderWeele et al, Cancer Science, 2014). This work fulfills the following subtasks from the SOW: subtask 2.1, subtask 2.3, subtask 3.1.

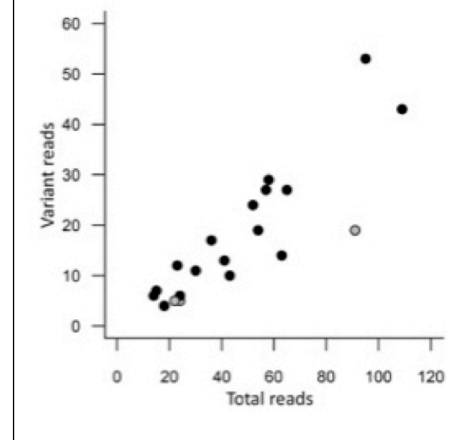
Aim 1b: *Determine if coincident acinar and ductal adenocarcinoma are clonally related.*

Prior to submitting the award proposal, I had obtained approval for IRB protocol 10-270-A to obtain deidentified prostate tissue specimens for use in this project. Shortly after the PRTA was awarded, this protocol was amended to allow use of samples from the PCBN. The PCBN was able to provide a single specimen of coincident acinar and ductal adenocarcinoma with sufficient material to prepare sequencing libraries.

I have worked with the Human Tissue Resources Core at the University of Chicago to identify additional cases of coincident ductal and acinar carcinoma with sufficient material to prepare sequencing libraries. Not all the specimens were of sufficient purity for macrodissection, and so a number have undergone laser microdissection (Figure 3). I aimed to evaluate at least 5 specimens. I have isolated acinar and ductal tumor DNA from 6 specimens in addition to the one from the PCBN, with an additional 4 cases available in case I again run into issues of sample quality as I did for Aim 1a.

This work addresses the following subtasks from the SOW: Subtask 1.1, subtask 1.3. I had planned on completing sequencing library preparation for these samples (included in subtask 1.3). That has not been completed yet but will be started once I have obtained normal DNA from adjacent tissue. I expect obtaining normal tissue (to be completed soon) to be far simpler than the task of performing laser microdissection to obtain tumor tissue (already completed). A minor setback was that I was expecting to be able to use more samples from the PCBN, but not as many were available as I had thought. I have made up for that with samples from the University of Chicago.

Figure 1. Confirmation of high confidence somatic mutations with capillary sequencing by number of total and variant reads (gray = not confirmed; black = confirmed). Confirmation rate = 0.85



At the time the proposal for this project was submitted, next-generation sequencing costs were decreasing rapidly, and thus I proposed doing whole genome sequencing of these samples. However, sequencing costs have since plateaued and whole genome sequencing is not as affordable as was anticipated. Therefore, instead of deep coverage whole genome sequencing, I plan to perform deep coverage exome sequencing to identify single nucleotide variants, and low coverage whole genome sequencing to provide information about copy number changes and identify potential structural variants. This strategy has been employed by many others, including The Cancer Genome Atlas project, with great success (REF###). The copy number changes and structural variants will allow evaluation of clonal relationships at a genome-wide level, which is what deep coverage whole genome sequencing was intended to do.

Aim 2: Characterize the molecular relationship between metastatic disease and primary prostate foci.

Aim 2a: *Test the hypothesis that nodal metastases are subclones of high grade prostate cancer foci in the prostate.*

As in Aim 1a, I have completed sequencing and data analysis of tumor DNA from lymph node metastasis from two patients with coincident low and high grade foci in the prostate (Table 1). A subset of the high confidence somatic mutations identified by VarScan have been validated by Sanger sequencing and custom capture and resequencing (Figure 1).

The results of this sub-Aim confirm that nodal metastases are closely related to high grade prostate cancer, and distantly related to low grade prostate cancer (Figure 2). Notably, the data demonstrate that

Figure 2. Venn diagrams (left) depicting the pattern of shared and private high confidence somatic mutations and phylogenetic trees (right) depicting the relationship of coincident foci for PrCa 18 (A), PrCa 14 (B), PrCa 6 (C), and PrCa 25 (D). The number of high confidence somatic mutations is labeled within each Venn diagram. In C, there are no mutations that are shared between the low grade focus and the metastatic focus and not also shared with the high grade focus. Gray = uninvolved prostate; blue = low grade focus; green = high grade focus; purple = metastatic focus; black = theoretical common progenitor.

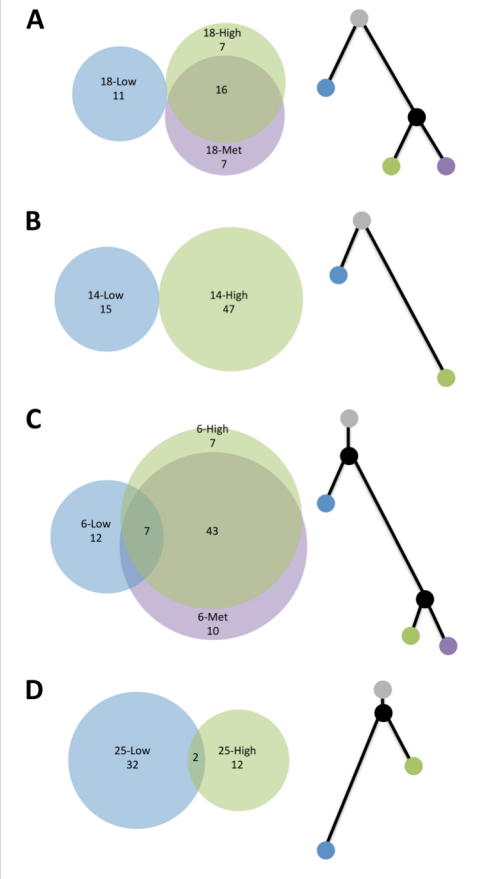
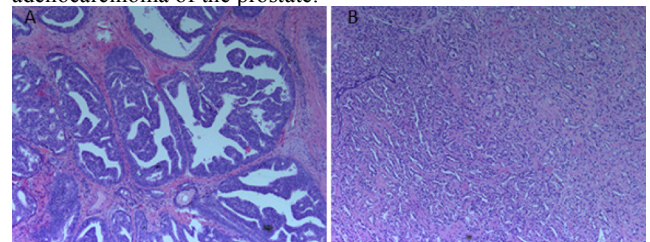
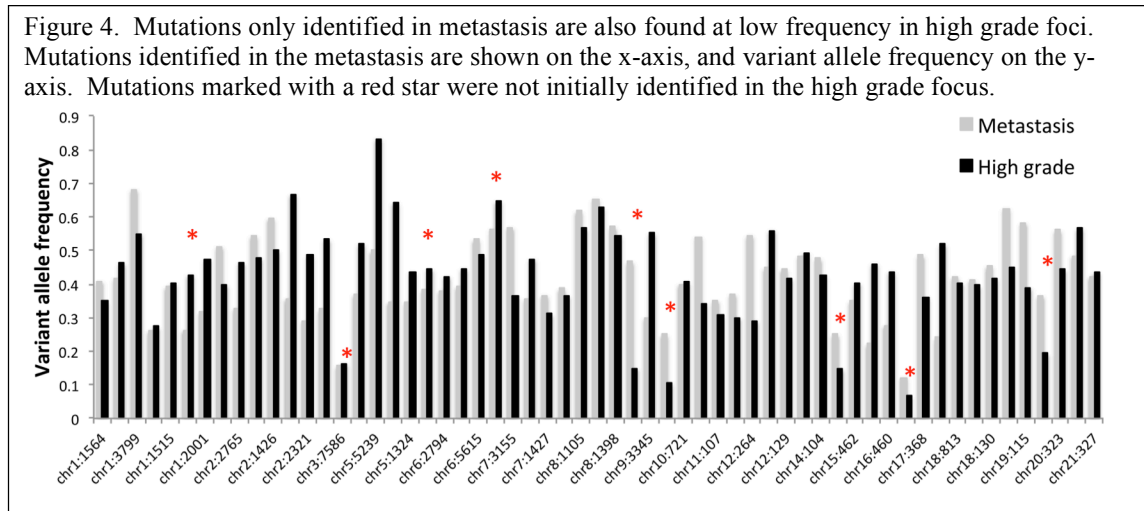


Figure 3. H & E stain of ductal (A) and acinar (B) adenocarcinoma of the prostate.



few mutations are required for metastases beyond what is found in high grade cancer. The data are also consistent with the hypothesis that nodal metastases are subclones of high grade prostate foci, for mutations identified in lymph node metastases but not initially identified in high grade foci can be identified at low levels in high grade foci (Figure 4).

Results from this aim are now published and the manuscript is included as an appendix to this report (VanderWeele et al, Cancer Science, 2014). This work fulfills the following subtasks from the SOW: subtask 2.1, subtask 2.3, subtask 3.1.



Aim 2b: *Determine the molecular relationship of CTCs to primary prostate cancer foci.*

To address this aim, I have obtained DNA isolated from CTCs from 7 patients, with biological replicates from three of those patients. However, library preparation and sequencing have not yet begun. There are two tasks related to this aim: evaluating the feasibility of sequencing CTCs, which involves sequencing biological replicates from two subjects; and determining the relationship of CTCs to primary prostate cancer foci. Given the financial cost of conducting these experiments, along with my mentors I decided it is better to consolidate these efforts. Furthermore, in addition to the relationship between CTCs and primary cancer foci, I am also interested in the relationship between CTCs and synchronous metastases to determine if CTCs represent current metastatic disease. Therefore I have switched to recruiting exclusively patients for whom primary tissue and metastatic tissue are available. We have found numerous patients from whom we've collected CTCs, who do not have primary or metastatic tissue available, and a publication evaluating protein expression in those patients' CTCs is accepted for publication. However, it has been more challenging to find patients with both these tissue specimens available.

Nevertheless, two patients have been recruited with primary and late metastatic disease, both of whom underwent CTC collection twice. One has biological replicate CTC DNA that has been whole genome amplified. CTC DNA from the 2nd patient failed to amplify on one of the replicates, and he is scheduled to undergo CTC collection again within a few weeks. To avoid batch effects I have delayed sequencing library

preparation until all the samples from both patients can be prepared at the same time. I anticipate that will begin within 1-2 months. In collaboration with Dr. Szmulewitz I have identified 7 additional patients with metastatic disease and primary tissue available who will be recruited for CTC collection when they are progressing on their current therapies, 3 of whom are scheduled to be collected within the next month.

This work addresses subtask 1.2 and subtask 1.4. Work on subtask 2.2 is expected to occur within 3 months. Though there has been a delay in performing the sequencing, I believe that the sequencing data we obtain will be more valuable because I am selecting patients with more informative specimens available.

As in Aim 1b, at the time the proposal for this project was submitted, next-generation sequencing costs were decreasing rapidly, and thus I proposed doing whole genome sequencing of these samples. Due to persistent high costs, however, instead of deep coverage whole genome sequencing, I plan to perform deep coverage exome sequencing to identify single nucleotide variants, and low coverage whole genome sequencing to provide information about copy number changes and identify potential structural variants.

Mentorship

Dr. Walter Stadler continues to be my mentor for clinical and translation research and has continued to promote my professional development. I meet with him formally monthly and informally several times throughout the week.

Dr. Yusuke Nakamura, Professor of Medicine, is my new primary basic science mentor. He is one of the pioneers of applying the study of genetic variation to the medical field. He joined the faculty at University of Chicago in 2012 in the department of Medicine and continues his research in cancer genomics, especially that of prostate cancer and other GU cancers. He is a recent recruit to the University of Chicago and his interests align more closely with my interests did those of Dr. Kevin White. In addition, as a physician and member of the section of hematology/oncology he is more focused on translational aspects of his work. I meet with Dr. Nakamura weekly on an informal basis and as part of the Nakamura group lab meetings. I meet with him more formally monthly.

Dr. Funmi Olopade continues as a mentor as an expert in cancer genetics and genetics of hormone-regulated cancers. She is also an important advisor regarding career development.

Key Research Accomplishments

- Publication of the first manuscript from this project, “Low grade prostate cancer diverges early from high grade and metastatic disease.”

Conclusion

Like other cancers, prostate cancer is a genetic disease that progresses through evolution and accumulation of genetic changes. Though low grade and high grade prostate cancer can evolve from a common progenitor, they are distantly related, and in some cases arise independently. In contrast, metastatic disease is closely related to high

grade localized disease. Certain genomic alterations likely mark high grade, high risk disease, and these includes changes in the TP53 pathway.

Publications, Abstracts and Presentations

1. VanderWeele DJ, Brown CB, Taxy JB, Stadler WM, White KP. Low grade prostate cancer foci diverge early from high grade and metastatic disease. *Cancer Science*. 2014 Aug;105(8):1079-85.
2. VanderWeele, DJ, et al. "Feasibility of multimodal molecular characterization of the androgen receptor (AR) in circulating tumor cells (CTCs) from patients with castrate-resistant prostate cancer (CRPC)." *JOURNAL OF CLINICAL ONCOLOGY*. AMER SOC CLINICAL ONCOLOGY, 2014.
3. Reyes E*, VanderWeele DJ*, Duggan R, Campanile A, Stadler WM, Vander Griend DJ, Szmulewitz R. Quantitative characterization of androgen receptor expression in circulating tumor cells from patients with metastatic castrate resistant prostate cancer. *Journal of Translational Medicine*. In press.

Inventions, Patents and Licenses

Nothing to report

Reportable Outcomes

Nothing to report

Other Achievements

The data from this work, and the protected time afforded by this award, led to my successful application for two additional grants and another pending application which passed initial evaluation and is now under review. These include a pilot award from our institution's cancer center, a Young Investigator Award from the Cancer Research Foundation, and an award from Psi Beta Phi. The work supported by this award has led to me being recruited as a faculty member at University of Chicago, and I am also a member of the University of Chicago Comprehensive Cancer Center.

I have successfully transitioned to having my own lab, under the mentorship of the three excellent mentors involved with this award. I have a research professional and a technician in the lab, last year I mentored a urology resident working in the lab, and this past summer I mentored an undergraduate student who was part of the Chicago Academic Medicine Program at the Pritzker School of Medicine.

I continue to be interested in prostate cancer tumor heterogeneity, and in order to obtain higher quality clinical specimens I have established a tissue bank of prostatectomy specimens and biopsy specimens fixed in the non-crosslinking agent PAXgene. I've also collected matched blood, plasma, and urine samples.

Opportunities for Training

As per the SOW, I have attended the following courses:

From year 1:

Responsible Conduct of Research

Ecology and Evolution: Principles of Population Genetics I

Ecology and Evolution: Genomic Evolution

From year 2:

Human Genetics: Introduction to Probability and Statistics for Geneticists

From year 3:

I am currently taking Statistics: Statistical Methods and their Applications

The Ecology and Evolution: Fundamentals of Molecular Evolution course I had planned for year 1 is no longer offered and there is no equivalent course in its place.

Opportunities for Professional Development

I have had the opportunity to attend and participate in many conferences during the first year of this award. I presented a poster on my work looking at low and high grade cancer foci at the Prostate Cancer Foundation Scientific Retreat. This meeting brings together the brightest minds in prostate cancer research along with representatives from industry and policy.

I presented a poster on our initial work with CTCs at the GU cancer symposium, which is a more clinically oriented meeting affiliated with the American Society of Clinical Oncology. I also attended the ASCO annual meeting.

I am the institutional PI for several clinical trials, including a trial evaluating neoadjuvant chemotherapy for high risk prostate cancer, which is run through the CALGB/Alliance clinical trials consortium. As my role I also participate in the semi-annual CALGB/Alliance meetings.

References

Appendices

VanderWeele DJ, Brown CB, Taxy JB, Stadler WM, White KP. Low grade prostate cancer foci diverge early from high grade and metastatic disease. Cancer Science. 2014 Aug;105(8):1079-85.

Low-grade prostate cancer diverges early from high grade and metastatic disease

David J. VanderWeele,^{1,2} Christopher D. Brown,^{1,3,6} Jerome B. Taxy,^{4,7} Marc Gillard,² David M. Hatcher,⁵ Westin R. Tom,⁵ Walter M. Stadler² and Kevin P. White^{1,2,3}

¹Institute for Genomics and Systems Biology, Departments of ²Medicine, ³Human Genetics, ⁴Pathology, ⁵Surgery, University of Chicago, Chicago, Illinois, USA

Key words

Carcinoma/genetics, exomes, genetic heterogeneity, multiple primary neoplasms, prostatic neoplasms

Correspondence

David VanderWeele, Department of Medicine, Section of Hematology/Oncology, The University of Chicago, 5841 South Maryland Avenue MC2115, Chicago, Illinois 60637, USA.
Tel: 773-834-7193; Fax: 773-702-9268;
E-mail: dvanderw@medicine.bsd.uchicago.edu

⁶Present address: Department of Genetics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA

⁷Present address: Department of Pathology and Laboratory Medicine, NorthShore University Health System, Evanston, Illinois, USA

Funding information

This study is supported by the AACI Fellowship for Translational Cancer Research and DOD Prostate Cancer Research Program PRTA (DVW) and the Chicago Cancer Genomes Project (KPW).

Received January 8, 2014; Revised May 16, 2014; Accepted May 23, 2014

Cancer Sci 105 (2014) 1079–1085

doi: 10.1111/cas.12460

Cancers are thought to progress from low grade to higher grade,⁽¹⁾ leading to efforts to detect and eradicate low-grade disease in order to prevent morbidity and mortality from high-grade disease. The grade of prostate cancer is characterized by the Gleason score on a scale of 2–10, with 10 representing the most poorly differentiated tumors. Although initially described as intermediate grade, Gleason 6 cancers are clinically indolent and have an overwhelmingly favorable prognosis.^(2,3) For the purposes of this report, Gleason 6 foci are regarded as low grade.

Approximately 238 590 men were diagnosed with prostate cancer in the US in 2013, the majority of them with low-grade disease. It is estimated that 29 720 men will die from prostate cancer, which is among the lowest case/fatality ratio of any cancer.⁽⁴⁾ Although prostate specific antigen (PSA) screening is able to identify early stage prostate cancer, most men diagnosed with prostate cancer die from competing causes. It is therefore unclear whether early detection associated with PSA screening decreases mortality.^(5,6) Current clinical treatment guidelines support active surveillance for patients with

Understanding the developmental relationship between indolent and aggressive tumors is central to understanding disease progression and making treatment decisions. For example, most men diagnosed with prostate cancer have clinically indolent disease and die from other causes. Overtreatment of prostate cancer remains a concern. Here we use laser microdissection followed by exome sequencing of low- and high-grade prostate cancer foci from four subjects, and metastatic disease from two of those subjects, to evaluate the molecular relationship of coincident cancer foci. Seventy of 79 (87%) high-confidence somatic mutations in low-grade disease were private to low-grade foci. In contrast, high-grade foci and metastases harbored many of the same mutations. In cases in which there was a metastatic focus, 15 of 80 (19%) high-confidence somatic mutations in high-grade foci were private. Seven of the 80 (9%) were shared with low-grade foci and 65 (82%) were shared with metastatic foci. Notably, mutations in cancer-associated genes and the p53 signaling pathway were found exclusively in high-grade foci and metastases. The pattern of mutations is consistent with early divergence between low- and high-grade foci and late divergence between high-grade foci and metastases. These data provide insights into the development of high-grade and metastatic prostate cancer.

low-grade prostate cancer for which the risk of tumor progression is low.⁽⁷⁾ Recently some guidelines have even recommended against PSA screening.⁽⁸⁾ Failure to follow these guidelines is likely due, in part, to concern that low-grade foci progress to higher grade and metastatic disease.

Several studies have contributed to a growing catalog of genetic variation in prostate cancer,^(9–14) giving insight into single nucleotide variations, rearrangements and copy number variation prevalent in this disease. Aside from common rearrangements involving the ETS family of transcription factors, especially v-ets avian erythroblastosis virus E26 oncogene homolog (ERG), the mutational spectrum of prostate cancers is diverse.⁽¹⁰⁾ Although increasing attention is being paid to the overall molecular heterogeneity of cancers within a single patient, there is limited data in this regard specifically for prostate cancer. This is especially relevant for prostate cancer, which is often multifocal.

More specifically, it is not clear if prostate cancer develops as a multifocal disease in which the foci are clonally related and share the vast majority of somatic mutations, or if it is a

multicentric disease, with each focus having its own set of private somatic mutations and each representing an independent primary lesion. Recent studies have examined multifocal disease and found evidence of a monoclonal origin of coincident foci,⁽¹⁵⁾ even when the foci have different Gleason patterns.^(16,17) This conclusion is somewhat surprising, given the difference in clinical behavior of different Gleason patterns.

The present study targets multifocal disease, specifically with low-grade (Gleason 6) and high-grade (Gleason 8 or higher) foci, and compares the molecular relationship of these foci with each other and with synchronous metastatic disease. To expand the number of specimens available for the present study, prostatectomy samples that had been formalin fixed and paraffin embedded (FFPE) were used, with modifications made to standard protocols to increase the yield of material used for sequencing. Exome sequencing was used to achieve a comprehensive evaluation of all genes harboring somatic point mutations. The findings indicate that there is early divergence or complete independence of low-grade and high-grade disease, but late divergence of high-grade disease and synchronous metastases.

Materials and Methods

Tissue and laser capture microdissection (LMD). Paraffin-embedded prostate and lymph node tissue that was 1–5 years old and matched FFPE and frozen normal prostate tissue was obtained from the Human Tissue Resource Center (HTRC) at the University of Chicago with the approval of the Internal Review Board. Tissue was reviewed by a genitourinary pathologist (J.B.T.) and regions of interest were marked for further evaluation. Four specimens had sufficient low- and high-grade disease for LMD. Two specimens (PrCa 6 and PrCa 18) also had sufficient lymph node metastatic disease. Tumor foci and histologically normal prostate were microdissected from 10- μ m sections using a Leica Laser Microdissection (Wetzlar, Germany) system.

Exome library preparation. Genomic FFPE DNA was isolated using a QIAamp DNA FFPE Tissue Kit (Qiagen, Valencia, CA, USA) following the manufacturer's instructions, but increasing the incubation in Proteinase K to overnight and reducing the 90°C incubation to 30 min. DNA from frozen tissue was isolated using a QIAamp Mini Kit (Qiagen). DNA

was fragmented using a Covaris S2 sonicator (Woburn, MA, USA), with settings of 10% duty cycle, intensity 4, 200 cycles per burst and 120 s. The fragmented DNA was blunt-end repaired using T4 Polymerase, Klenow and T4 PNK, and adenylated. Fragments were ligated to adaptors IS1_adapter_P5.F and IS3_adapter_P5 + P7.R⁽¹⁸⁾ and an index tag added. The product was amplified and cleaned up using Agencourt AM-Pure XP beads (Beckman Coulter, Indianapolis, IN, USA).

Libraries were run on a BioAnalyzer (Agilent, Santa Clara, CA, USA) to confirm fragment size. The libraries from the metastases did not undergo further size selection. The remaining libraries underwent selection using an E-gel (Invitrogen, Grand Island, NY, USA) followed by PCR amplification.

Enrichment for exomes was performed using NimbleGen SeqCap EZExome v2.0 (Roche, Madison, WI, USA) following the manufacturer's instructions. The exome-enriched library DNA was amplified and run on a BioAnalyzer to confirm fragment size. The resulting libraries were run on a HiSeq 2000 (Illumina, San Diego, CA, USA) in a highly multiplexed 1 \times 50 run and a subsequent 2 \times 100 run.

Sequencing analysis and identification of somatic mutations. Overlapping or adapter-containing reads were clipped using SeqPrep-b83fd00. Reads were quality trimmed and aligned to hg18 using bwa 0.5.9⁽¹⁹⁾ and merged with duplicates removed using Samtools 0.1.18.⁽²⁰⁾ Genome Analysis Toolkit⁽²¹⁾ was used for local realignment and quality score recalibration. Somatic mutations were identified using VarScan v2.2.8.⁽²²⁾ High-confidence somatic mutations were defined as variants with coverage of at least 10 \times , no supporting reads in normal, at least four supporting reads in the tumor and with supporting reads on both strands. Germline single-nucleotide polymorphisms (SNPs) identified in dbSNP135, 1000 genomes and exomes of normal controls published by the Max Planck Institute (MPI)^(23–25) were removed. SAMtools was used to determine the number of total and variant reads in matched libraries for positions mutated in low-grade foci. Sample ethnicity was determined by principle component analysis using common, germline, coding variants shared with samples from the 1000 genomes project.⁽²⁴⁾

Custom capture and identification of previously identified mutations. DNA from seven additional foci from PrCa 6 were isolated using LMD (Leica Microsystems LMD6500). Custom capture probes targeting high-confidence mutations

Table 1. Subject characteristics

Case	Patient age (years)	Race	Stage	Gleason – gland†	Library	Gleason – focus‡
PrCa 18	65	Caucasian	pT3bN1Mx	4 + 3	Normal	NA
					Low grade	6
					High grade	8
					Metastasis	NA
PrCa 14	54	African–American	pT3bN1Mx	4 + 3	Normal	NA
					Low grade	6
					High grade	8
PrCa 6	70	Caucasian	pT3bN1Mx	4 + 3 (5)	Normal	NA
					Low grade	6
					High grade	8
					Metastasis	NA
PrCa 25	59	Asian	pT2cN0Mx	3 + 4	Normal	NA
					Low grade	6
					High grade	9

†Gleason score assigned on pathological review of the prostatectomy specimen. ‡Gleason score of the focus isolated by microdissection. PrCa 6 had a Gleason score of 4 + 3 with a minor component of 5. PrCa 14 had metastatic disease in a lymph node, which was not of sufficient mass to be isolated for sequencing. NA, not applicable.

identified through exome sequencing were obtained from Roche. Sequencing libraries from the seven loci were prepared using the High-Throughput Genome Analysis Core at the University of Chicago, pooled, captured following the manufacturer's instructions, and sequenced on one lane of a HiSeq2000 (Illumina). Reads were analyzed with SeqPrep-b83fd00, aligned to hg18 using bwa, and merged with duplicates removed using Samtools 0.1.18. Samtools was used to determine the number of total and variant reads in each library for each variant position. Variants were considered present if >10% of reads supported the variant.

Confirmation of somatic mutations and immunohistochemistry (IHC). Mutations were visualized using IGV 2.0.⁽²⁶⁾ Genomic DNA from each focus underwent amplification using GenomePlex WGA2 (Sigma, St. Louis, MO, USA) and primers targeting a subset of variants were designed using Primer3Plus and obtained from IDT (Coralville, IA, USA). The PCR products were sequenced in the University of Chicago DNA Sequencing and Genotyping Facility. ERG IHC was performed at the

HTRC using anti-ERG antibody EPR3864 (Abcam, Cambridge, MA, USA).

Results

To evaluate the molecular relationship between low- and high-grade prostate cancer, we identified multifocal formalin-fixed, paraffin-embedded specimens with both low-grade and high-grade disease with sufficient mass for laser capture microdissection. The subjects from whom the specimens came were representative of patients with high-risk localized prostate cancer (Table 1).

Exome sequencing was performed on matched cancer foci and histologically normal prostate glands from four specimens (Fig. 1). To improve the yield of DNA for downstream sequencing, proteinase K digestion was increased to 18 h, 90° incubation was reduced to 30 min and size selection by gel extraction was replaced with size selection with an E-gel or eliminated altogether. On average, 88% of RefSeq coding

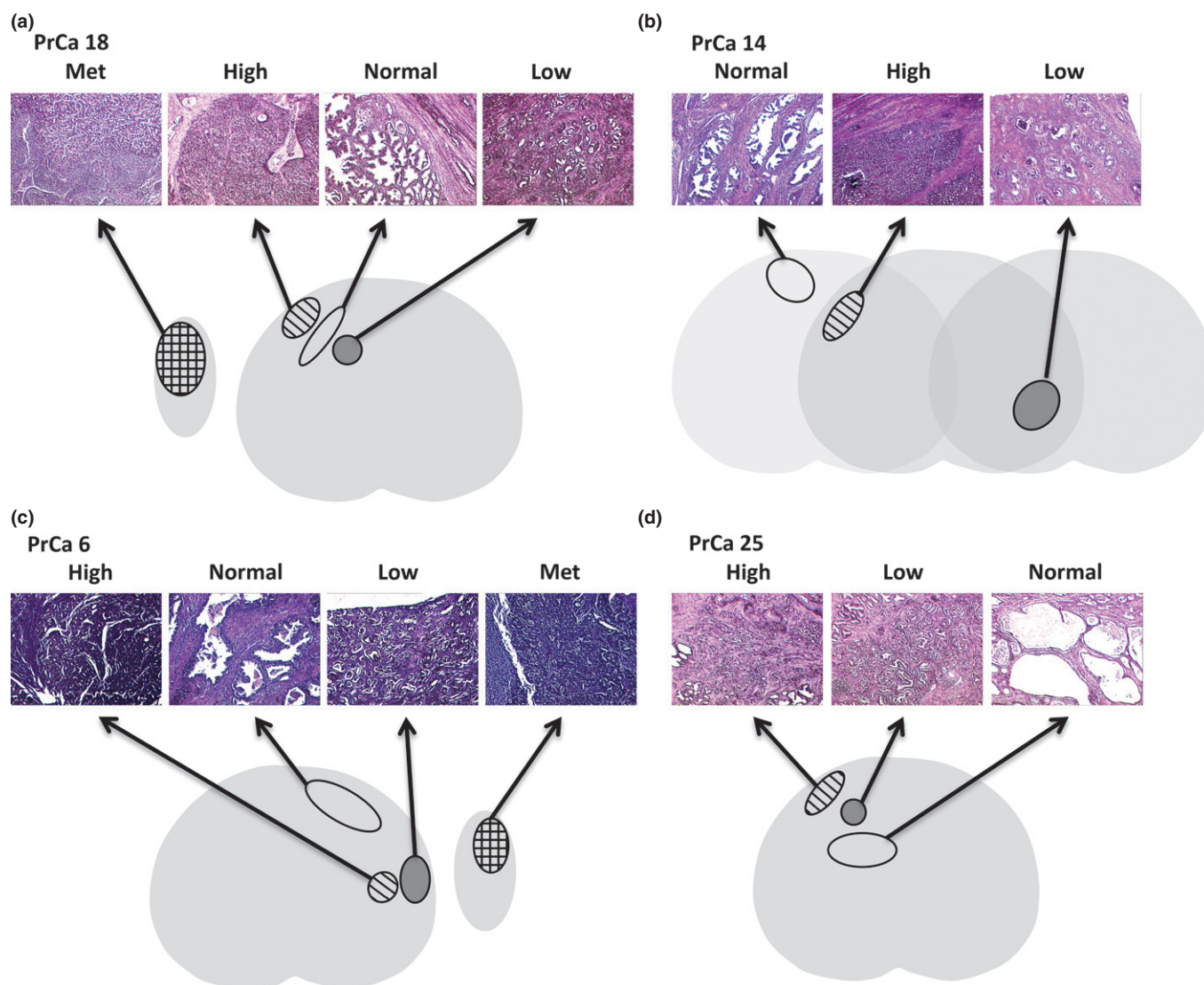


Fig. 1. Histology of coincident prostate cancer foci. Representative H&E stains and illustrations representing the prostate cancer foci that were laser microdissected. Two cancer foci and uninvolved prostate glands were isolated from PrCa 18 (a), PrCa 14 (b), PrCa 6 (c) and PrCa 25 (d). In addition, a metastatic (Met) focus was isolated from PrCa 18 and PrCa 6. For PrCa 14 (b), the foci were from different levels of the prostate. For the other three specimens the foci were at the same level. Light gray, histologically normal prostate; dark gray, low-grade cancer focus; striped, high-grade cancer focus; checked, metastatic focus from a lymph node removed at the time of prostatectomy.

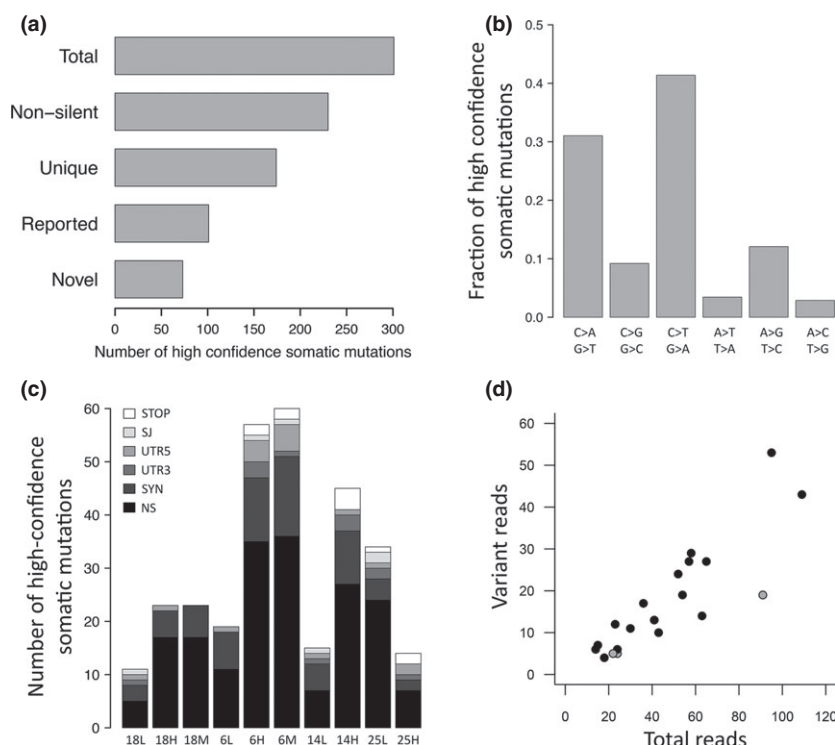


Fig. 2. Mutation characteristics. (a) Number of high-confidence somatic mutations across all foci. Non-silent, non-silent mutations; Unique, number of unique genes harboring a non-silent mutation; Reported, gene reported to be mutated in references 9–12 and 14. (b) Spectrum of unique high confidence somatic mutations across all foci. (c) Number of high-confidence somatic mutation types in all prostate samples. L, low-grade focus; H, high-grade focus; M, metastatic focus; NS, nonsynonymous; SYN, synonymous; UTR3, 3' untranslated region; UTR5, 5' untranslated region; SJ, splice junction; STOP, premature stop codon. (d) Confirmation of high-confidence somatic mutations with capillary sequencing by number of total and variant reads (gray, not confirmed; black, confirmed). Confirmation rate, 0.85.

bases were covered at 10× or greater. A mean of 30.1 somatic coding mutations was identified per cancer focus (Supporting Information Table S1).

To evaluate potential bias introduced by formalin fixation, matched FFPE and frozen (Optimal Cutting Temperature compound [OCT]-embedded) samples of normal prostate were subjected to exome sequencing. The number of false positive variant reads was similar (Fig. S1A), as were the types of mutations (Fig. S1B), suggesting that there was little bias introduced from formalin fixation.

Three hundred and one somatic mutations were identified across 10 foci (Table S2). Of these, 71 were synonymous and unlikely to alter gene function (Fig. 2a). Of 174 unique genes harboring somatic mutations that were likely to be functional, 73 were not reported in five previous reports of mutations in localized and advanced disease.^(9–12,14) The majority of variants identified were transitions leading to nonsynonymous mutations (Fig. 2b,c), as expected. A subset of 20 high-confidence somatic mutations was confirmed in tumor and uninvolved tissue with capillary sequencing (Fig. 2d). Those that were not confirmed by capillary sequencing had low variant allele frequency and were approximately the level of detection for capillary sequencing (~20%). Nevertheless, the mutation (0.82/Mb) and validation rate (85%) using fixed prostate specimens in the present study are similar to previous reports (0.9/Mb and 91%, respectively) using frozen tissue.⁽⁹⁾ The high fidelity of this sequencing supports the use of formalin-fixed tissue for next-generation sequencing studies.

The vast majority of somatic mutations in low- and high-grade foci were private to the focus in which they were identified. Nine of 79 (11%) high-confidence somatic mutations in the low-grade foci were shared with a high-grade focus, and nine of 139 (7%) high-confidence somatic mutations in the high-grade foci were shared with a low-grade focus (Fig. 3). Capillary sequencing confirmed that these nine shared somatic mutations were not germline variants. In two subjects (14 and 18) all of the mutations were private. In the other two subjects (6 and 25) a small

minority of mutations was shared. In contrast to the low-grade foci, 67 of 80 (84%) high-confidence somatic mutations identified in high-grade foci were shared with metastatic disease and metastatic foci had few mutations beyond those identified in high-grade disease (Fig. 3a,c).

To evaluate more broadly the clonal relationship of multifocal disease, custom capture baits targeting variants identified through exome sequencing of PrCa 6 were used to capture the DNA from seven additional cancer foci from that subject (Table S3). The average coverage depth of these variants was 208 and 78 previously identified variants were covered at >70× in each focus. Three high-grade foci harbored a median of 56 previously identified variants, three intermediate-grade foci harbored a median of 15 variants and one low-grade focus harbored three variants. As in the whole exome data a small minority of mutations was shared, with just two of 78 (2.6%) variants found to be ubiquitous among the seven cancer foci. Fourteen of 78 (17.9%) were not identified in any of the seven additional foci (Fig. S2). Eight of these 14 were among the 12 variants originally identified as private to the low-grade focus.

Rearrangements involving ERG are the most common genetic derangement in prostate cancer and are thought to occur early in tumorigenesis. These events are not typically captured by exome sequencing. Therefore we performed IHC for ERG to look for overexpression, which is typically restricted to cells harboring an ERG rearrangement. Of the two specimens exhibiting positive ERG staining, PrCa 6 showed consistently positive staining in all tumor foci (Fig. 4). In contrast, in PrCa 18 the low-grade focus was negative, the high-grade focus was equivocal with few cells positive for ERG and the metastatic focus was more widely focally positive.

Together with the immunohistochemical analysis, the patterns of shared and private mutations indicate early divergence of low- and high-grade foci. In two cases no shared mutations were identified and independent origin of low- and high-grade disease cannot be ruled out. In contrast, there is late divergence of high-grade foci and metastatic disease. Consistent

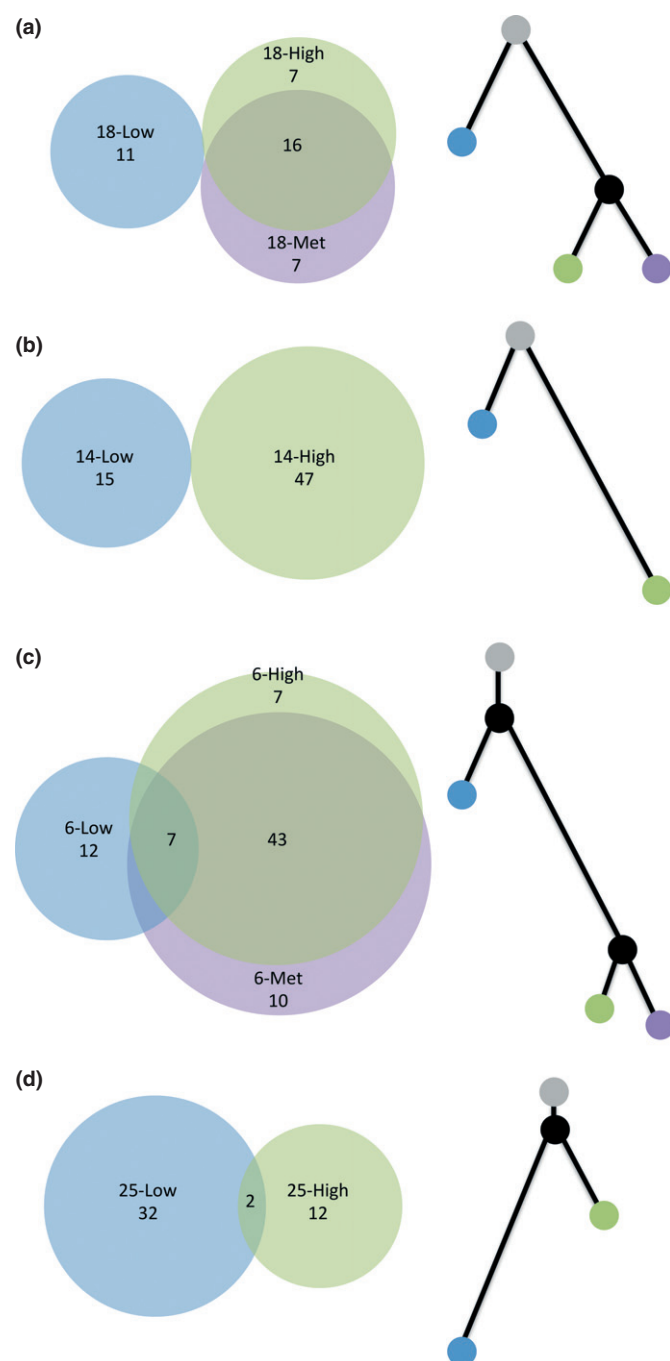


Fig. 3. The molecular relationship of coincident foci. Venn diagrams (left) depicting the pattern of shared and private high-confidence somatic mutations and phylogenetic trees (right) depicting the relationship of coincident foci for PrCa 18 (a), PrCa 14 (b), PrCa 6 (c) and PrCa 25 (d). The number of high-confidence somatic mutations is labeled within each Venn diagram. In (c), there are no mutations that are shared between the low-grade focus and the metastatic focus and also not shared with the high-grade focus. Within each phylogenetic tree, branch length is proportional to the number of mutations. Gray, uninvolved prostate; blue, low-grade focus; green, high-grade focus; purple, metastatic focus; black, theoretical common progenitor.

with this, there were no somatic mutations that were shared between low grade and metastatic disease and also not shared with high-grade disease.

Prostate tumor foci consist of an admixture of tumor cells and stromal cells. Despite microdissection of the tumor foci, the var-

iant allele frequencies in the low-grade foci were modestly lower than the high-grade lesions (Fig. S3), suggesting higher contamination with non-tumor cells. The contaminating stromal cells may interfere with the ability to call somatic mutations, especially in the low-grade foci, leading to a falsely elevated number of apparently private mutations and obscuring the clonal relationship among concurrent foci. Therefore for each position identified as a private high-confidence somatic mutation in the low-grade foci, the number of sequencing reads supporting that mutation in the matched high grade and metastatic foci was determined, relaxing filters based on DNA strand and quality scores (Fig. S4). Of the 23 high-confidence somatic mutations in PrCa 18 and PrCa 6 identified as private mutations, none had two or more variant reads in high grade or metastatic foci. Thus, even with a less conservative definition of mutation (two variant reads), high-grade foci and metastases harbored few variants seen in low-grade foci and the majority of variants in the low-grade foci was private. This was corroborated with evidence from capillary sequencing. Three mutations private to low-grade foci and two shared with other foci were among the subset of mutations confirmed by capillary sequencing.

Two genes exhibited recurrent nonsynonymous mutations or mutations in the untranslated region (Table S4). Low density lipoprotein (LDL) receptor-related protein 1B, which encodes an LDL-family receptor, is frequently deleted in multiple malignancies.⁽²⁷⁾ However, recent work suggests its high frequency of mutation may not imply a significant role in cancer.⁽²⁸⁾ ZNF717 is one of a number of zinc finger proteins and its role in malignancy is unclear.

Comparison of genes with nonsynonymous mutations in our data to the Catalogue Of Somatic Mutations In Cancer (COSMIC) cancer gene census revealed three genes in common (Table 2). These genes include TP53, the most frequently mutated gene in advanced cancers including prostate cancer,⁽¹²⁾ as well as ATM and SS18L1. All of these three genes were mutated in high grade or metastatic disease, but not in low-grade disease. Gene ontology analysis⁽²⁹⁾ revealed enrichment for genes in the p53 signaling pathway among high-grade foci (adjusted *P*-value 0.037), including TP53 and ATM. Cell cycle genes were also enriched (adjusted *P*-value 0.054). No other pathways were statistically significantly enriched.

Discussion

Given the disparate outcomes between men with low-grade prostate cancer and those with high-grade disease, a fundamental question in prostate cancer management is the possible molecular relationship among grades and metastatic disease. Multiple lines of evidence suggest multifocal disease is often independent.^(30,31) However, recent reports have demonstrated shared molecular derangements of coincident low- and high-grade foci.^(16,17) The present study identifies mutations shared between low- and high-grade cancer foci, supporting the ability of a single progenitor to give rise to both low- and high-grade disease. However, the relationship between these foci was distant relative to the relationship between high-grade disease and synchronous metastases. In fact, in two of the four cases it cannot be ruled out that the foci arose from independent origins. The early divergence of low- and high-grade disease may help explain distinct differences in their clinical behavior. Indeed, we found no evidence of direct progression from low grade to metastatic disease. Instead, there was an overwhelming fraction of shared mutations between high grade and metastatic disease, which is consistent with late divergence of these high-grade foci (Fig. 5).

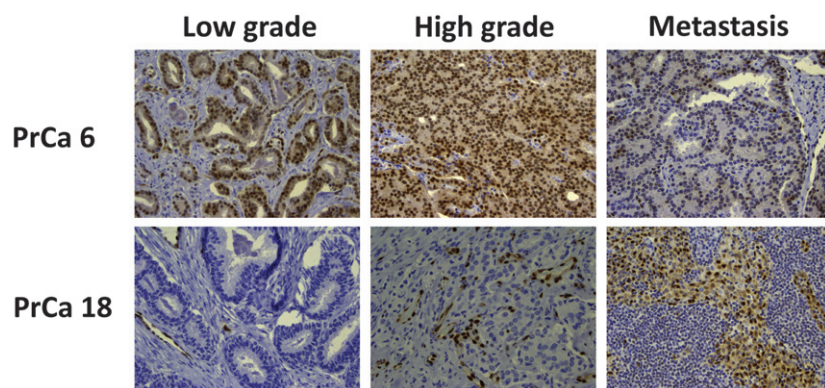


Fig. 4. ERG expression. Immunohistochemical analysis for ERG in coincident foci from PrCa 6 and PrCa 18.

Table 2. Overlap of mutated genes with the COSMIC cancer gene census

Subject	Histology	Chromosome	Hg18 position	Reference	Variant	Symbol	GERP	Type	aa Change
6	High	11	107675775	G	A	ATM	5.22	STOP	W1710*
6	LN	11	107675775	G	A	ATM	5.22	STOP	W1710*
6	High	20	60169915	C	G	SS18L1	3.98	NS	P87R
6	LN	20	60169915	C	G	SS18L1	3.98	NS	P87R
14	High	17	7517833	C	A	TP53	4.78	NS	C145F

Three genes harboring a nonsynonymous mutation or a premature stop codon are also found in the COSMIC cancer gene census. All three were found in either high-grade foci or metastatic disease, but not in low-grade foci. GERP, genomic evolutionary rate profiling score; LN, lymph node metastasis; NS, nonsynonymous; STOP, introduction of premature stop codon; *, stop codon.

For the cases in which there was evidence of a common progenitor for both low- and high-grade disease (PrCa 25 and PrCa 6), the nature of the common progenitor is not clear. Histological low-grade cancer, HGPIN or even histologically uninvolved prostate are all possible progenitors. If the progenitor is not histological low-grade cancer, high-grade and low-grade disease may have emerged simultaneously. Characterization of the common progenitor requires further study.

The present study demonstrates there is late divergence of high-grade cancer and synchronous metastatic disease, confirming high-grade disease as the precursor to metastasis and suggesting few if any additional mutations are required for malignant cells to gain the ability to metastasize. Given the difficulty in obtaining tissue from typical bone metastases in prostate cancer, more easily accessible prostate lesions might offer a substitute for identifying genetic derangements leading to primary resistance of metastatic disease. However, mechanisms of secondary resistance are likely induced or enriched by the selective pressure of therapy and thus are likely found at low levels, if at all, in the primary lesion.

The most common gene fusion event in prostate cancer is the TMPRSS2-ERG fusion, leading to high levels of expression of ERG, which is otherwise expressed at very low levels in prostate cancer. We evaluated for the presence of this fusion through IHC for ERG and those findings supported the evidence from the exome sequencing data. A limitation of exome sequencing is the inability to identify additional chromosomal rearrangements. Although this is a limitation it was not required for the goal of the present study, which was to identify tumor lineages.

One goal of sequencing prostate tumors is to identify biomarkers that will supplement the prognosis given by clinical features alone. To that end we sought to identify mutated genes or pathways that were enriched in high grade and metastatic disease. TP53 is the most frequently mutated gene in advanced prostate cancer and we found the p53 signaling pathway to harbor mutations only in high grade and metastatic disease. Thus, members of the p53 signaling pathway are

intriguing candidates for potential biomarkers of aggressive disease. Recently, expression of a panel of cell cycle progression-associated genes has been correlated with outcomes of those with low- or intermediate-grade prostate cancer.⁽³²⁾ We also found enrichment for cell cycle-related genes among those mutated in high-grade disease, although this was of borderline statistical significance in this small cohort.

Given the good outcomes of low-grade prostate cancer, we hypothesized that mutations not found in low-grade disease are more likely to be drivers of progression and metastasis and therefore more important potential therapeutic targets. Indeed, the three genes identified here that are also found in the COSMIC cancer gene census were identified exclusively in high grade and metastatic foci. Of note, all three mutations are in highly conserved regions as evidenced by the high genomic evolutionary rate profiling score and thus are likely to be biologically significant. Identification of driver genes through this strategy has the potential to guide the development of future therapies.

The present study contributes to the growing body of work demonstrating that next-generation sequencing from microdissected FFPE prostatectomy specimens is feasible. Based on these techniques, we found that coincident low- and high-grade prostate cancer can emerge through either independent progression or early divergence from a common progenitor.

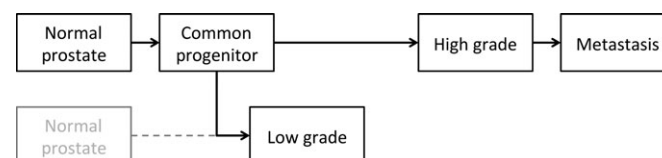


Fig. 5. Model of prostate cancer evolution. Low-grade and high-grade cancer foci progress largely in parallel, diverging early from a common progenitor. Metastatic disease exhibits late divergence from high-grade disease. It cannot be ruled out that in some cases low- and high-grade diseases arise through independent origins.

High-grade disease is characterized by mutations in known cancer-associated genes and the p53 signaling pathway. Metastatic disease is closely related to high-grade primary foci. These data can be leveraged to help identify potential biomarkers and drivers of progression of clinically significant disease.

Acknowledgements

The authors thank April Peterson for technical assistance in library preparation and A. Jason Grundstad, Dr Megan McNerney and Dr Thomas Stricker for advice on experimental design and data analysis.

The present study is supported by the AACI Fellowship for Translational Cancer Research and DOD Prostate Cancer Research Program PRTA (D.J.V.W.) and the Chicago Cancer Genomes Project (K.P.W.).

Disclosure Statement

The authors have no conflicts of interest.

References

- Vogelstein B, Kinzler KW. The multistep nature of cancer. *Trends Genet* 1993; **9**: 138–41.
- Egger SE, Scardino PT, Walsh PC *et al.* Predicting 15-year prostate cancer specific mortality after radical prostatectomy. *J Urol* 2011; **185**: 869–75.
- Ross HM, Kryvenko ON, Cowan JE, Wheeler TM, Epstein JI. Do adenocarcinomas of the prostate with Gleason score (GS) ≤ 6 have the potential to metastasize to lymph nodes? *Am J Surg Pathol* 2012; **36**: 1346–52.
- Siegel R, Naishadham D, Jemal A. Cancer statistics, 2013. *CA Cancer J Clin* 2013; **63**: 11–30.
- Andriole GL, Crawford ED, Grubb RL 3rd *et al.* Mortality results from a randomized prostate-cancer screening trial. *N Engl J Med* 2009; **360**: 1310–9.
- Schröder FH, Hugosson J, Roobol MJ *et al.* Prostate-cancer mortality at 11 years of follow-up. *N Engl J Med* 2012; **366**: 981–90.
- Mohler J, Bahnson RR, Boston B *et al.* NCCN clinical practice guidelines in oncology: prostate cancer. *J Natl Compr Canc Netw* 2010; **8**: 162–200.
- Moyer VA. Screening for prostate cancer: U.S. Preventive Services Task Force recommendation statement. *Ann Intern Med* 2012; **157**: 120–34.
- Berger MF, Lawrence MS, Demicheli F *et al.* The genomic complexity of primary human prostate cancer. *Nature* 2011; **470**: 214–20.
- Barbieri CE, Baca SC, Lawrence MS *et al.* Exome sequencing identifies recurrent SPOP, FOXA1 and MED12 mutations in prostate cancer. *Nat Genet* 2012; **44**: 685–9.
- Grasso CS, Wu Y-M, Robinson DR *et al.* The mutational landscape of lethal castration-resistant prostate cancer. *Nature* 2012; **487**: 239–43.
- Kumar A, White TA, MacKenzie AP *et al.* Exome sequencing identifies a spectrum of mutation frequencies in advanced and lethal prostate cancers. *Proc Natl Acad Sci USA* 2011; **108**: 17087–92.
- Taylor BS, Schultz N, Hieronymus H *et al.* Integrative genomic profiling of human prostate cancer. *Cancer Cell* 2010; **18**: 11–22.
- Robbins CM, Tembe WA, Baker A *et al.* Copy number and targeted mutational analysis reveals novel somatic events in metastatic prostate tumors. *Genome Res* 2011; **21**: 47–55.
- Boyd LK, Mao X, Xue L *et al.* High-resolution genome-wide copy-number analysis suggests a monoclonal origin of multifocal prostate cancer. *Genes Chromosomes Cancer* 2012; **51**: 579–89.
- Sowalsky AG, Ye H, Bubley GJ, Balk SP. Clonal progression of prostate cancers from Gleason grade 3 to grade 4. *Cancer Res* 2013; **73**: 1050–5.
- Kovtun IV, Chevillat JC, Murphy SJ *et al.* Lineage relationship of Gleason patterns in Gleason score 7 prostate cancer. *Cancer Res* 2013; **73**: 3275–84.
- Meyer M, Kircher M. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harb Protoc* 2010; **2010**: pdb.prot5448.
- Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 2009; **25**: 1754–60.
- Li H, Handsaker B, Wysoker A *et al.* The sequence alignment/map format and SAMtools. *Bioinformatics* 2009; **25**: 2078–9.
- McKenna A, Hanna M, Banks E *et al.* The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 2010; **20**: 1297–303.
- Koboldt DC, Zhang Q, Larson DE *et al.* VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res* 2012; **22**: 568–76.
- Sherry ST, Ward MH, Kholodov M *et al.* dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res* 2001; **29**: 308–11.
- Consortium T 1000 GP. A map of human genome variation from population-scale sequencing. *Nature* 2010; **467**: 1061–73.
- Burbano HA, Hodges E, Green RE *et al.* Targeted investigation of the Neanderthal genome by array-based sequence capture. *Science* 2010; **328**: 723–5.
- Robinson JT, Thorvaldsdóttir H, Winckler W *et al.* Integrative genomics viewer. *Nat Biotechnol* 2011; **29**: 24–6.
- Prazeres H, Torres J, Rodrigues F *et al.* Chromosomal, epigenetic and microRNA-mediated inactivation of LRP1B, a modulator of the extracellular environment of thyroid cancer cells. *Oncogene* 2011; **30**: 1302–17.
- Lawrence MS, Stojanov P, Polak P *et al.* Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* 2013; **499**: 214–8.
- Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 2008; **4**: 44–57.
- Mehra R, Han B, Tomlins SA *et al.* Heterogeneity of TMPRSS2 gene rearrangements in multifocal prostate adenocarcinoma: molecular evidence for an independent group of diseases. *Cancer Res* 2007; **67**: 7991–5.
- Lindberg J, Klevebring D, Liu W *et al.* Exome sequencing of prostate cancer supports the hypothesis of independent tumour origins. *Eur Urol* 2013; **63**: 347–53.
- Cooperberg MR, Simko JP, Cowan JE *et al.* Validation of a cell-cycle progression gene panel to improve risk stratification in a contemporary prostatectomy cohort. *J Clin Oncol* 2013; **31**: 1428–34.

Supporting Information

Additional supporting information may be found in the online version of this article:

Fig. S1. Comparison of fixed and frozen.

Fig. S2. Variant allele frequency of variants in additional cancer foci.

Fig. S3. Variant allele frequencies.

Fig. S4. Frequencies of variants identified as private to the low-grade focus.

Table S1. Characteristics of sequencing libraries.

Table S2. High-confidence somatic mutations identified across all foci sequenced.

Table S3. Variants identified in additional cancer foci.

Table S4. Recurrently mutated genes.